Learning symbolic features for rule induction in computer aided diagnosis

Sebastijan Dumančić[1,2], Antoine Adam[1] and Hendrik Blockeel[1]

[1] Department of Computer Science, Katholieke Universiteit Leuven, 3001 Heverlee, Belgium
sebastijan.dumancic@student.kuleuven.be
[2] Faculty of Electrical Engineering and Computing, University of Zagreb, 10 000 Zagreb, Croatia

Abstract. In computer aided medical diagnosis (CAD), interpretability of learned models is an important concern. Unfortunately, the raw data used to train a model are often in subsymbolic form (for instance, images), which makes the application of symbolic learning methods difficult. Construction of symbolic features can bridge the gap between the symbolic and subsymbolic level. This paper presents a case study of how ILP learners can be used to learn models from visual data by using a feature construction step. The resulting model has an accuracy comparable to that of previous models, but better interpretability.

Keywords: computer aided diagnostics, inductive logic programming, deep learning

1 Introduction

Ever more frequently, and in an increasing variety of domains, machine learning based models assist humans in decision making. The accuracy of learned models often matches or even exceeds that of human experts. Despite this, learned models may not be used because they are "black boxes": their decisions are based on computations not understandable to humans. In some domains, interpretability of the model is essential.

An example of such a domain is computer aided diagnosis (CAD) in medicine. Consider the "anti-nuclear antibodies" (ANA) test, which is used to diagnose autoimmune diseases. The diagnosis is based on visually identifying certain staining patterns in cells. Example patterns are shown in Figure 1. Currently, the test is not fully automated; a physician, looking at an image, has to decide what staining pattern occurs. This decision is known to be subjective; it depends heavily on the expertise of the physician, and on the varieties of reading systems and optics. Recently, black box algorithms have achieved very good accuracy on this task [5]. However, unless the automated system can explain its decision, physicians are unlikely to blindly trust it, given the subjectiveness of the decision process.

In this paper, we want to break open the black box. We propose to learn interpretable models from raw image data by introducing a feature construction step that extracts symbolic features using subsymbolic learning. The system we present here consists of three steps. First, individual cells are segmented from the images obtained by Indirect immunofluorescence imaging. Second, each segmented cell is mapped to several symbolic features describing their visual properties. Third, an ILP system learns an interpretable model on these extracted features. To our knowledge, this is the first system (for this application domain) that covers the whole path from image to decision (including segmentation), and also the first that yields interpretable models.

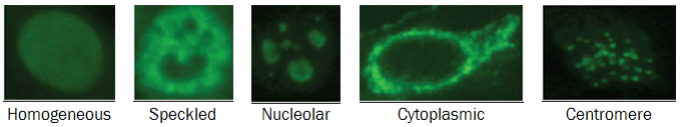

Fig. 1. Examples of HEp-2 staining patterns

2 Proposed approach

The approach is illustrated in figure 2. The system we present consists of three steps. First, individual cells are segmented from the images obtained by Indirect immunofluorescence imaging. Second, each segmented cell is then mapped to several symbolic features describing their visual appearance. This mapping is partially hardcoded and partially learned using SVMs or deep learning. Finally, we use an ILP system to learn an interpretable model on these extracted features.

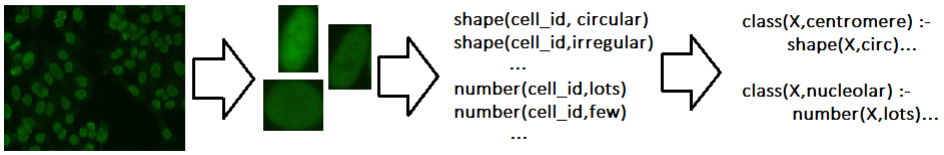

Fig. 2. Illustration of a process

2.1 Segmentation

Our segmentation procedure first determines the background color; it is easily found as a narrow peak in the image histogram. A region-growing algorithm next

finds a large connected region with this color; this is considered to be the background. Next, a standard method from image recognition, the Hough transform [7], is used to identify roughly circular objects formed by non-background pixels. Finding roughly circular objects has the advantage that when two or more cells touch, the individual cells may still be identified. The Hough transform results in circles around the identified objects; these circles are evolved to tight-fitting shapes using Morphological Snakes [8]. These shapes are our segmented objects.

2.2 Symbolic feature learning

This next step is the key to the interpretability of our final model. In this step, we want to extract features that will be used as an input by the classifier in the last step. This could be seen as a preprocessing step. However, we do not want any features to be extracted. We want features that are understandable by physicians. Therefore, we extract features that are used in the medical literature. These features and their possible values are the following:

- **shape**: circular, irregular
- **fluorescence intensity level**: positive, intermediate
- **structure**: homogeneous, speckled
- **organelle type**: dark, bright, neutral
- **organelle number**: none, few, lots
- **texture**: smooth, sparkly, blob
- **mitotic cell type**: bright middle, dark middle, neutral, speckled (for now, identified manually)

Most of these features describe visual properties of cells. We train Deep Belief Networks [10] to extract these, as they are known to work well for such tasks. One network per feature is learned, with the feature values as classes. This is done in two steps - first unsupervised by training Restricted Boltzmann machines, then supervised using backpropagation [12].

Two features are learned in a different way: the shape and the fluorescence intensity level. For the shape, we only want to identify if it is circular or not, and efficient computer vision algorithm have been designed for shape recognition. We used a simplified method of shape contexts by Belongie et al [17]. Unlike other features, the fluorescence intensity level does not describe properties of the cell itself but how visible the cell is. Based on the assumption that, in grey scale, the image consists of bright cells and a dark background, we fit a mixture of two Gaussians to the image histogram and train an SVM [9] on the parameters of these Gaussians with rbf kernels.

2.3 Rule induction

The symbolic representation obtained in the previous step is the crucial part for building an interpretable model. The second choice that has to be made is the

classifier that supports such representation. Since interpretability is crucial, tree or rule learners or ILP systems are natural choices for this step.

3 Experiments

For our experiments, we use the HEp-2 dataset published for the International Conference on Pattern Recognition 2012.[3] This dataset contains raw images as well as separate individual cells extracted from them. We have manually annotated the cells with values for the features discussed above. This gives a dataset that has one instance per cell, and each instance is described using features the values of which have been determined by a human; we consider these values the "ground truth". Call this dataset TRUE. We have also derived a second dataset from the raw images by performing automatic segmentation, annotating each cell with feature values as predicted by models trained on TRUE. Call this dataset SYS. Finally, we have learned classification rules from TRUE and SYS. Note that, in a sense, the rules learned from TRUE are learned using exact values for the input features, while the SYS rules can be considered to have noisy values.

The description of the cells is currently output in first-order logic format, and hence ILP systems are a natural choice for the learner to be used. We have experimented with FOIL [14] and Aleph [15]. As our current representation is not inherently relational, it is also possible to translate it to a propositional format, run a propositional learner, and translate the rules back to clauses.; we have used the RIPPER [16] system in this way. Among these approaches, RIPPER turned out to obtain the highest accuracy as well as the most compact models.

Table 1 shows how accurately the features can be predicted using our learned models. Table 2 shows the performance of RIPPER, in terms of recall and precision, for the TRUE and SYS rules. In terms of accuracy, the TRUE rules achieve 94.5% accuracy (measured using tenfold cross-validation), and the SYS rules achieve 88.7%. The best black-box system until now achieved 95.59 %. This shows that, while the input features are predicted with relatively good accuracy, the noise introduced by this prediction still reduces the accuracy of the final prediction; if correctly extracted, however, the features would allow for rule-based prediction on par with the best black-box methods. Figure 3 shows some examples of induced rules. The complete model learned from TRUE contains 10 rules and 13 literals; the one learned from SYS contains 12 rules and 16 literals.

Table 1. Performance of symbolic feature learning

Feature	shape	intensity	structure	number of organelles	organelle type	texture
Accuracy	98.68 %	96.21 %	91.4 %	89.7 %	93.78 %	90.22 %

[3] `http://mivia.unisa.it/hep2contest/index.shtml`

Table 2. Performance of the RIPPER classifier

		homogeneous	nucleolar	centromere	cytoplasmatic	fine speckled	coarse speckled
TRUE	Prec	93.48 %	94.65 %	98.24 %	96.33 %	83.33 %	92.78 %
	Rec	100 %	95.44 %	93.54 %	96.33 %	86.54 %	85.71 %
SYS	Prec	92.94 %	86.40 %	94.67 %	98.02 %	81.40 %	90.91 %
	Rec	96.70 %	91.95 %	89.64 %	90.83 %	81.13 %	80.95 %

```
class(X, cytoplasmatic) :- shape(X, irregular)
class(X, cytoplasmatic) :- texture(X, blob)
class(X, centromere) :- organelles_type(X, bright), organelle_number(X, lots)
class(X, homogeneous) :- not organelle_number(X, lots), mitotic_cells(X, bright_middle)
```

Fig. 3. Examples of induced rules

4 Conclusion

Construction of interpretable models from subsymbolic data can be a non-trivial task. In this paper, we have proposed a method that does this in the context of computer aided diagnosis (CAD). The method consists of (1) defining symbolic features that are interpretable to humans; (2) learning a first layer of models that map raw data (images) onto a description based on these features; (3) learning an interpretable model that maps the feature-based description to some target variable. Experiments show that, on the domain considered, this interpretable model can achieve accuracy comparable to black-box models.

We used a supervised approach to learn features, but one can think of using unsupervised learning to define the features; this would effectively amount to "feature invention", or "monadic predicate invention" in ILP terms. Although the features defined here are just properties of single objects, and in this application we have just one type of object (cells), one can think of learning properties of different types of objects in this way, and learning to combine these on a higher level using a relational learner. This would make it possible to learn relational theories from subsymbolic data. Finally, it is an open question to what extent "relational features" could be learned, or non-monadic predicates invented; this seems a very challenging task.

Within this particular application, other possible future work includes automatic mitotic cell detection and artefact removal, broader experimentation with deep learning approaches and introduction of uncertainty in rules.

Acknowledgements

This work is funded by the KU Leuven Research Fund (project IDO/10/012).

References

1. Foggia, P. et al.: Benchmarking HEp-2 Cells Classification Methods. IEEE Trans. Med. Imaging, vol. 32, 1878-1889 (2013)
2. Huang, Y.L. et al.: Adaptive Automatic Segmentation of HEp-2 Cells in Indirect Immunofluorescence Images. IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, IEEE Computer Society, 418–422 (2008)
3. Huang, Y.L. et al.: Outline Detection for the HEp-2 Cell in Indirect Immunofluorescence Images Using Watershed Segmentation. IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, IEEE Computer Society, pp. 423–427 (2008)
4. Agrawal,P., Vatsa, M., Singh, R.: HEp-2 Cell Image Classification: A Comparative Analysis. Lecture Notes in Computer Science, Machine Learning in Medical Imaging, vol. 8184, Springer International Publishing, 195-202, (2013)
5. Williem, A. et al: Classification of Human Epithelial Type 2 Cell Indirect Immunofluoresence Images via Codebook Based Descriptors. 2013 IEEE Workshop on Applications of Computer Vision (WACV), IEEE Computer Society, 95–102, (2013)
6. Dempster, A. P., Laird, N. M., Rubin,D. B.: Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, series B, vol. 39, pp. 1–38 (1977)
7. Hough, P. V. C.: Machine Analysis of Bubble Chamber Pictures. International Conference on High Energy Accelerators and Instrumentation, vol. C590914, 554-558 (1959)
8. Marquez-Neila, P., Baumela, L., Alvarez, L.: A Morphological Approach to Curvature-Based Evolution of Curves and Surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, IEEE Computer Society, 2–17 (2014)
9. Vapnik, V., Cortes, C.: Support-Vector Networks. Machine learning, vol. 20, Kluwer Academic Publishers, 273–297 (1995)
10. Bengio, Y.: Learning Deep Architectures for AI. Found. Trends Mach. Learn., vol. 2, Now Publishers Inc., 1–127 (2009)
11. Larochelle, H., Bengio,Y.: Classification using discriminative restricted boltzmann machines. 25th international conference on Machine learning, ACM (2008)
12. Murphy, K. : Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series). MIT Press (2012)
13. Lavrač, N. and Džeroski,S. : Inductive Logic Programming: Techniques and Applications. Routledge (1993)
14. Quinlan, J. R. earning logical definitions from relations Machine Learning, vol. 5, 239–266 (1990)
15. Muggleton, S. Inverse Entailment and Progol New Generation Computing, vol 13, 245–286 (1995)
16. Cohen, William W. Fast Effective Rule Induction Proceedings of the Twelfth International Conference on Machine Learning, Morgan Kaufmann, 115–123 (1995)
17. Belongie, S., Malik, J. and Puzicha, J. Matching Shapes International Conference on Computer Vision, vol. 1, 454–461 (2001)
18. Palm, R. B. : Prediction as a candidate for learning deep hierarchical models of data. Master's Thesis (2012)